SEEDS

Working Paper Six:

ESE Logical Data Service Provider Types

April 24, 2002

G. Hunolt, SGT, Inc.

Outline

1 Introduction

This working paper is the sixth of a set of papers that describes the SEEDS (Strategic Evolution of Earth Science Enterprise Data Systems) Levels of Service (LOS) / Cost Estimation (LOS/CE) study. The study goal is to develop a cost estimation model and coupled requirements and levels of services to support the SEEDS Formulation team in estimating the life cycle costs of future ESE data service providers and supporting systems, where 'data service provider' is used as a generic term for any data/information related activity. The set of working papers is intended to serve as a vehicle for coordinating work on the project, obtaining feedback and guidance from ESDIS SOO and the user community, and as embryos of reports that will be produced as the task proceeds.

As working papers, each version of each paper that appears represents a snapshot in time, with the work in various stages of completion. As work progresses the content (and sometimes the organization) of the working papers will change reflecting progress made, responses to feedback and guidance received, etc.

This sixth working paper of the set describes the current set of logical data service provider types developed for the LOS/CE study, and reflects results of the February, 2002, SEEDS Community Workshop.

Section 2 discusses the logical data service providers as subsets of the general data services provider model.

Section 3 presents descriptions of the logical data service provider types.

Section 4 presents a mapping of the reference model requirements and levels of service from Working Paper 5, "Data Service Provider Reference Model - Requirements / Levels of Service" to the logical data service provider types in Section 3.

2 Logical Data Service Provider Types in Relation to the Data Service Provider Reference Model

The Data Service Provider Reference Model is a general functional model of an abstract, generic data service provider that includes a full set of functional areas. The model is described by the set of functional areas, corresponding requirements and levels of service, and a parameter set that is mapped to the functional areas and levels of service. (See Working Paper 3, "Data Service Provider Reference Model - Functional Areas", Working Paper 4, "Data Service Provider Reference Model - Model Parameters", and Working Paper 5, "Data Service Provider Reference Model - Requirements and Levels of Service".)

2.1 Logical Data Service Provider Types as Reference Model Subsets

The general data service provider reference model includes all functions / areas of cost that a generic data service provider might perform. While an actual working data service provider could conceivably perform all of the functions included in the model, most if not all actual data service providers perform a subset of them, e.g. most providers will not have a requirement in the area of instrument / mission operations. Many well known actual data centers such as the NASA Distributed Active Archive Centers or the NOAA national data centers perform a subset of the general set of functions. Some data service providers, e.g. MODAPS (the MODIS Adaptive Processing System, a sample of a science team processing facility that does not perform archive or general user distribution), are different in function from many well known data centers but fit within the framework of the data service provider reference model.

The Cost Estimation Tool will allow a planner, for example one planning a data service to support a flight project, to:

1) select those functions that are required for his/her particular mission (in effect creating a 'custom' subset of the model);

2) specify the particular mission requirements the real instantiation of it must meet (e.g. data volumes to be ingested, processed, stored, and/or distributed);

3) produce an estimated cost for implementing and operating it.

To facilitate overall ESE data service architecture studies (where a 'data service architecture' is a collection of data service providers and the interconnections between them), a set of 'logical data service provider types' has been defined. Each of these types is a functional subset of the general reference model organized around a defined class of ESE role or mission. These are 'logical' types in that there is no explicit or implicit 1:1 mapping of an instance of a logical data service provider type to a physical entity. While some actual data service providers might match a logical type, most will perform the functions of more than one logical type, and may also perform multiple data service activities within the scope of a type (such as a DAAC that performs archive and distribution for several flight projects). Because the logical data service provider types are only a few of the possible subsets of the general model, they constitute an open set to which additions (and subtractions) can be readily made as needed to facilitate architecture trade studies or other uses.

2.2 Using Logical Data Service Provider Types

The logical data service provider types can be used in two ways, as discussed in Working Paper 2, "Cost Estimation by Analogy Model", which includes scenarios showing how the Cost Estimation Tool would be used. Two different modes of use of the tool are described.

The first mode is to produce a life-cycle cost estimate for a particular data service provider activity to be performed by a new provider or as an additional task by an existing provider. In this case the user of the tool would select the data service provider functions needed in the particular case, and produce an estimate of the cost for implementing and operating it. For this purpose, while a user would have the freedom to create a

custom set of functions (in effect creating a custom subset of the general data service provider model) the user would also have the option of deciding that his/her needs corresponded to a logical data service provider type and using a template for it to facilitate producing the cost estimate.

The second mode is to produce an overall estimate for an ESE architecture, some combination of organizations performing data services functions such that the aggregate ESE requirements for data services are met. As described in Working Paper 2, producing a cost estimate for an overall ESE architecture, or a number of different estimates for alternative architectures, requires the user to deal with a large number of data service activities, having to encompass the ESE as a whole. The user in this situation will not be able to consider any one activity exhaustively, and for this purpose the logical data service provider types will be of great help.

In either case, the logical types must be useful subsets of the general reference model, i.e. they must be organized around an ESE role or mission that is significant in the real world. Only then will they be of genuine value to the ESE data services architect, or attractive to the individual or team planning a single activity.

3   ESE Logical Data Service Provider Types

This section describes the current set of ESE logical data service provider types, drawing on the NewDISS concept paper "Draft Version 1.0 - NewDISS: A 6-to-10-year Approach to Data Systems and Services for NASA's Earth Science Enterprise", October 2000 for a starting point. For each logical data service provider type, this section will present the conceptual description taken from the concept paper and a description of the functions of the data service provider type in terms of the data service provider reference model and its functional areas (defining the subset of the reference model that applies to the data service provider type).

The NewDISS concept paper introduces its discussion of NewDISS data service provider types: "NASA's ESE has requirements for collection and synthesis of scientific information, for bringing synthesized data products to bear on unanswered scientific questions, and for preserving data and information for future scientific discovery. … NewDISS is therefore seen as consisting of a dynamic network of interconnected components, each responsive to its environment, containing capabilities for change over time through feedback with the science community. These components will be responsible for executing NewDISS data management functions and must allow easy participation by scientists and data and services providers. The components of NewDISS have been conceptualized (October, 2000) as including "Backbone" processing centers, PI-managed Mission Data Centers, Science Data Centers, and Multi-Mission Centers [here Systematic Measurement Centers]."

Three additional data service provider types are added:

1. Applications Center, focused on uses and users other than research, given the existence of NASA funded applications activities such as Type III Earth Science Information Partners (ESIPs) and Regional Earth Science Applications Centers (RESACs);

2. Information Center, focused on information describing data and products rather than the data and products themselves, based on discussion at the Formulation Team Retreat, November 7-8, 2001, where 'Echo' was suggested as a possible future instance, and the Global Change Master Directory (GCMD) is plainly a currently operational instance;

3. Long Term Archive Center, focused on permanent preservation and archiving of data and products and their documentation and active support to climate research, etc., based on a request from Matt Schwaller, a member of the Formulation Team and leader of the Earth Science Data Life-Cycle study. Long term archiving is strictly speaking not an ESE responsibility, but inclusion of a hypothetical Long Term Archive data service provider type is intended to support planning that NASA is doing with NOAA and USGS, the agencies who have (with the National Archives and Records Administration, NARA) the long term archive responsibility.

A "data service provider" does not necessarily imply a physically distinct institution. An institution such as a NASA center, a university, an organization of another US Government Agency such as USGS or NOAA can host a data service provider or a combination of data service providers. This is equivalent to the existing situation in which the University of Colorado hosts the National Snow and Ice Data Center (NSIDC) DAAC, or the USGS's EROS Data Center (EDC) hosts the EDC DAAC.

3.1   Backbone Data Center

This section describes the generic Backbone Data Center type.

3.1.1   Backbone Data Center Concept

The following is the concept for Backbone Data Centers, from the NewDISS Concept Paper: "These centers, most likely evolving from some of the current DAAC's, will address NASA's responsibility for preserving and protecting the large volumes of data from the ESE satellite missions. One of the primary roles of the backbone data centers will be to preserve the basic data. Clearly, NASA can provide a considerable amount of existing

infrastructure and technical skill needed to provide satellite mission data downlink and "level 0" or "level 1" data processing. Teaming NASA missions with Backbone Data Centers in the Announcement of Opportunity (AO) process for backup or for generation of basic data products may well be an attractive option for handling some of the core data management requirements of NewDISS. Another role for the Backbone Data Centers will be to acquire products agreed to be scientifically important for preservation and to prepare all these data for long-term archiving. These data centers will need to address network connectivity as part of their on-going activities. Selection of these services will be driven by PI-teaming arrangements, using either NASA-available resources or competitive alternatives. Backbone Data Centers, staffed by professional data managers, provide a core set of historical experience and proven capabilities. As such, they provide a means for risk mitigation against the failure of one or more of the NewDISS components by serving as backup centers for the other parts of the NewDISS. These data centers would most likely be few in number to ensure the cost-effectiveness of the NewDISS."

3.1.2 Backbone Data Center Functions

In general, the Backbone Data Center is expected to provide stable and highly robust services, with a key responsibility for data preservation and documentation, and with a mandate to provide professional data management as a resource for ESE as a whole. A Backbone Data Center is not identified with a particular mission or project but provides data management services in support of multiple missions and the NASA science program in general. The Backbone Data Center would have an indefinite lifespan subject to regular independent review of its performance.

In addition to its operational functions, the Backbone Data Center would bring a high level of expertise in the areas of data stewardship, science, and information technology in planning and conducting its own activities, working with its science user (and other as appropriate) community to understand how best to meet its needs, and coordinating with ESE and other ESE data service providers in many areas (standards, interfaces, interoperability, data interuse across providers, best practices, data stewardship, user support, information technology, etc.).

The Backbone Data Center would be expected to move to increasingly automated services, e.g. allowing user software to interact directly with center's system to provide the functional equivalent of a manual search and order system and direct access to data and products. Backbone Data Centers may also increasingly support data integration and data mining for users, using software developed by the center for the purpose or in some cases using software provided by the user.

The Backbone Data Center could provide processing functions for NASA missions through teaming arrangements with NASA Principal Investigators, and can serve as a backup to other ESE data service providers.

The paragraphs below will briefly discuss the Backbone Data Center role in each of the general data service provider reference model's functional areas.

Ingest - The Backbone Data Center performs ingest of a wide variety of data types, ranging from low level data streams to ancillary data to all of the levels of derived products, including their metadata, documentation, etc. In some cases the ingest function must be performed on a time critical, operational basis, e.g. for data and supporting information received from operating satellite platforms via NASA or other agency mission operations and communications systems. Level of service agreements or equivalent (e.g. operations agreements, interface agreements) with sources may be required. Quality control on incoming data is especially critical for lower level (e.g. level 0) data ingested, as the Backbone Data Center must detect bad data and request replacement data from operational sources that may have a limited capability for storing and retransmitting data.

Processing - The Backbone Data Center may perform processing through a teaming arrangement with a flight mission Principal Investigator, which can include large scale (in terms of number of products generated and /or product volume data) operational, schedule driven 'standard product' processing and reprocessing, perhaps

with emphasis on Level 1 processing vs higher level derived product processing. The Backbone Data Center would provide a science software integration and test service. Operational processing by the Backbone Data Center would be highly reliable with tight quality control.

The Backbone Data Center may also perform non-operational processing which could include generation of research products, data integration products, and data mining.

Documentation - The Backbone Data Center ensures that its data and product holdings are documented to the SEEDS adopted standard for long term archiving, working as necessary with external data sources (e.g. other data service providers) to capture all needed information. The Backbone Data Center also ensures produces search metadata, product guides, etc., to SEEDS adopted standards, which would likely include Federal Geographic Data Committee (FGDC) compliant metadata.

Archive - The Backbone Data Center provides a very robust archive capability, performing insertion of data into archive storage, and performing archive quality data stewardship, including preservation of data, metadata, and documentation within the archive. Preservation measures should include quality screening of data entering and exiting the archive, quality screening of archive media, off-site backup with sampling and tested restoration to verify integrity, and accomplishing migrations from one type of media to another.

Search and Order - The Backbone Data Center serves a broad user community with a robust and flexible search and order capability that supports user interaction with search and order services and, increasingly, supports automated search and order interaction between software running on a user system and the Backbone Data Center system. The search capability allows a user to apply criteria that might include geophysical parameter(s), spatial-temporal coverage, specific product names, etc., to the metadata describing available data and products and returning to the user listings supplemented by descriptive information of those data or product types and instances that meet the criteria. The 'order' capability includes a request/permission step, regardless of how implemented (e.g. manual or automated), where a request for a set of data or product instances, perhaps the results of (or a selected subset of the results of) a search, is processed and accepted or denied.

Backbone Data Center search and order can include providing local user interface and capability and/or providing an interface to a broader based, ESE cross-site search and order capability.

Access and Distribution - The Backbone Data Center serves a broad user community with a robust access and distribution (electronic and media) service, including offering data services such as subsetting, reformatting, reprojecting, packaging in response to user needs. The Backbone Data Center will increasingly support automated access to its data and products by user software.

The Backbone Data Center will also transfer data and documentation to designated long term archive centers in accordance with life cycle data management plans.

User Support - The Backbone Data Center provides effective user support for a wide range of users, including support provided in direct contact with users by user support staff, e.g. responding to queries, taking of orders, staffing a help desk (i.e., staff awaiting user contacts who can assist in ordering, track and status pending requests, resolve problems, etc.), etc. The Backbone Data Center will increasingly offer more automated user support aids (beginning with on-line documentation, FAQ, etc.) to meet increasing demands on user support with the proliferation of data types, data sources, and tools for users. User support staff should include science expertise to provide users with assistance in selecting and using data.

The Backbone Data Center coordinates its user support with other ESE data service providers (e.g. for user referral services). It performs outreach to potential new users, and participates in coordinated outreach activities with other ESE data service providers.

Instrument / Mission Operations - The Backbone Data Center does not perform this function.

Sustaining Engineering - The Backbone Data Center performs sustaining engineering, with no or very infrequent interruption of operational capabilities.

Engineering Support - The Backbone Data Center performs engineering support functions with no or very infrequent interruption of its operations.

Technical Coordination - The Backbone Data Center is heavily involved in technical coordination. It participates in SEEDS system level processes, including coordination on data management, documentation standards, data stewardship (including standards for content of life cycle data management plans), standards and best practices (including quality assurance standards and practices), interfaces, common metrics, and interoperability (e.g. for data access and integration), across / within SEEDS and with other systems and networks as needed to support the ESE program.

The Backbone Data Center participates in ongoing examination of the changing needs of the ESE science and applications program and the consequent impacts on the roles, missions, and services of ESE data service providers.

The Backbone Data Center participates in coordination of user support guidelines and practices across the network of ESE data service providers and with other data centers as needed to support the ESE science and applications program.

The Backbone Data Center cooperates with other ESE data service providers in representing ESE / SEEDS in broader community processes in areas such as standards, interoperability, data management, security, etc.

The Backbone Data Center also participates in SEEDS level and/or bilateral processes to coordinate production and delivery of products between itself and other ESE data service providers.

Implementation - The Backbone Data Center develops the data and information system capabilities it requires to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

The Backbone Data Center also maintains an ongoing applications software development effort. Applications software can include software to perform data services (e.g. subsetting, reformatting, reprojection, etc.) for more of its products, software tools for use by users to unpack, subset, or otherwise manipulate products provided by the Backbone Data Center, product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need, and to perform a 'data mining' or data integration operation to meet a user need.

Management - The Backbone Data Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Backbone Data Center includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. It includes developing data stewardship practices, performing data administration with science advice (via the User Advisory Group and other appropriate bodies), developing and maintaining life cycle data management plans (which address data migrations). It also includes coordinating its internal science activities and its interaction with the ESE and broader science community, including a visiting scientist program or equivalent, collaboration among ESE data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group (which includes representation from the science, applications, education, etc., communities that it serves) and any other ESE or broader advisory activities that may be appropriate.

The Backbone Data Center also participates in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS.

Facility / Infrastructure - The Backbone Data Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings, including a separate off-site backup facility for its data and information holdings. The Backbone Data Center ensures system and site security according to established NASA security policies and practices.

The Backbone Data Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

The Backbone Data Center provides facility / infrastructure support at a level that ensures no or very infrequent interruption of its operations.

3.2 Mission Data Center

This section describes the generic Mission Data Center type.

3.2.1 Mission Data Center Concept

The following is the concept for NewDISS Mission Data Centers, from the NewDISS Concept Paper: "These data systems are specifically affiliated with instruments or satellite systems. They are either PI-led or facility/project-led. They provide key measurements and standard products from NASA-supported satellite instruments. The key characteristic of the mission data centers is that they will be engineered and implemented as part of an ESE mission proposal. It is anticipated that these Mission Data Centers could leverage the activity at the current ESE data management infrastructure: the ECS flight operations and science data systems and the other hardware and software infrastructure at the DAAC's, the ESIP's, and the SCF's. These data centers will need to address network connectivity as part of their on-going activities. Selection of these services will be driven by PI-teaming arrangements, using either NASA-available resources or competitive alternatives. Mission Data Centers will also need to address satellite/instrument command and control and data downlink. Selection of these services will be driven by PI-teaming arrangements, using either NASA-available resources or competitive alternatives, such as commercially provided or university support services."

Mission Data Centers will be responsible for their data management functions during an Earth-observation space flight mission. These data service providers will be funded by the mission selected through the ESE flight programs and will be selected by competitive selection for future ESE missions."

3.2.2 Mission Data Center Functions

In general, the Mission Data Center is an element of a particular ESE mission that exists to provide data management services for the life of that mission. The mission might be might involve an instrument on an independently operated spacecraft (such as SeaWinds on a Japanese platform) or might include multiple instruments on a dedicated spacecraft (such as Terra or Aqua). The services provided by the Mission Data Center extend from instrument or platform command and control through generation and distribution to mission science team members of science products derived from instrument data for quality assurance, validation, and research. A Mission Data Center would provide instrument data and science products to a Backbone Data Center for distribution to the broad user community and archive after the mission life is completed.

Ingest - The Mission Data Center ingests instrument and spacecraft telemetry and instrument data from NASA or other spacecraft operations and communications systems, and ancillary data needed to support product generation from various sources. Ingest of instrument data and instrument and spacecraft telemetry might be performed on a time critical, operational basis, and the Mission Data Center must detect bad data and request replacement data from operational sources that may have a limited capability for storing and retransmitting data.

Processing - The Mission Data Centers will perform small to large scale (in terms of number of products generated and /or product volume data) 'standard product' processing and reprocessing. If the processing is performed to meet the needs of the mission science team only, it can be performed as the team requires. If the processing also must meet the needs of other missions (e.g. as ancillary products), science teams, or other users, it may be performed on an operational basis (especially once processing algorithms become stable). Processing by the Mission Data Center would include tight quality control. The Mission Data Center could

team with a Backbone Data Center for the processing service, especially if there is a requirement for routine, operational generation of 'standard' products.

Documentation - The Mission Data Center generates complete documentation of its instrument data and all derived products. The Mission Data Center cooperates with a Backbone Data Center that receives its data after completion of its mission to ensure that documentation is brought to long term archiving standards.

Archive - The Mission Data Center would not perform an archive function per se, but would maintain secure working storage of data and products until their transfer to a Backbone Data Center at some time during the mission or after completion of the mission. The Mission Data Center would maintain an off-site back up of all data for which it is responsible, and might use the services of a Backbone Data Center for this purpose.

Search and Order - The Mission Data Center serves its mission science team with a robust and flexible search and order capability tailored to meet the needs of the science team.

Access and Distribution - The Mission Data Center provides products to the mission science team for quality assurance, validation, or research, with a search and order capability as needed to meet the needs of the mission science team. The Mission Data Center will also transfer data, products, and documentation to a Backbone Data Center either during its mission as backup or when broader distribution of its data and products is appropriate, or at the conclusion of the mission.

User Support - The Mission Data Center provides close support to member of the mission science team.

Instrument / Mission Operations - The Mission Data Center performs this function for instruments and spacecraft that are part of its mission through NASA or other appropriate operational mission management services. This includes monitoring instrument and spacecraft performance, generating instrument and (if applicable) spacecraft commands, and event scheduling.

Sustaining Engineering - The Mission Data Center performs sustaining engineering, with no or very infrequent interruption of any critical operational capabilities.

Engineering Support - The Mission Data Center performs engineering support functions as needed, but with no or very infrequent interruption of any critical operational capability.

Technical Coordination - The Mission Data Center participates in certain SEEDS system level processes, including coordination on data management, documentation standards and best practices (including quality assurance standards and practices), interfaces, security, and common metrics.

The Mission Data Center also participates in SEEDS level and/or bilateral processes to coordinate production and delivery of products between itself and other ESE data service providers.

Implementation - The Mission Data Center develops the data and information system capabilities it requires to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement as needed over its mission life.

The Mission Data Center also maintains an ongoing applications software development effort. Applications software would include 'science software' - product generation software embodying science algorithms, e.g. to produce a suite of products to meet the needs of the mission's research program and the overall ESE research program. In some cases the science software would be developed to run in the Mission Data Center's own environment, in other cases the Mission Data Center could provide science software to a Backbone Data Center for operational production.

Management - The Mission Data Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Mission Data Center includes developing and maintaining life cycle data management plans for data generated by its mission, coordinating with other data service providers as needed,

e.g. a Backbone Data Center to which the mission data and products and complete documentation would be transferred to after the end of the mission.

It also includes coordinating its interaction with the ESE and broader science community, collaboration among ESE data service providers to support science needs, and any ESE or broader advisory activities that may be appropriate.

The Mission Data Center also participates in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS as needed for the success of its mission.

Facility / Infrastructure - The Mission Data Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and working storage for its data and information holdings, including a separate off-site backup facility, for which the Mission Data Center might use the services of a Backbone Data Center. The Mission Data Center ensures system and site security according to established NASA security policies and practices.

The Mission Data Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

The Mission Data Center provides facility / infrastructure support at a level that ensures no or very infrequent interruption of its operations.

3.3   Science Data Center

This section describes the generic Science Data Center type.

3.3.1   Science Data Center Concept

The following is the concept for NewDISS Science Data Center, from the NewDISS Concept Paper: "These data centers will collect data from multiple missions for a user community focused on a single research question. There are several examples of these types of Science Data Centers in NASA's Space Science Enterprise. These centers are targeted at specific science questions (perhaps from the NRC Pathways Report) and/or science disciplines, and they directly support research and data analysis for specific research questions. These data centers will address network connectivity as part of their on-going activities. Selection of these services will be driven by PI-teaming arrangements, using either NASA-available resources or competitive alternatives."

3.3.2   Science Data Center Functions

In general, the Science Data Center is a temporary data management capability implemented to support a particular research effort by a limited community of users (which will be called its 'research team'). A Science Data Center could support a 'data mission' - supporting a research team doing science with existing data without a new flight project. The research effort could be interdisciplinary or focused on one of the traditional Earth science disciplines. The Science Data Center operates in a research environment, without the same need for robustness and performance as would be the case for an operational environment.

Ingest - The Science Data Center obtains data and products required to meet the research objectives of its research team from a variety of sources, including other ESE data service providers, other agency data centers, etc. The ingest would not be performed on a time critical, operational basis.

Processing - The Science Data Center would perform non-operational processing, and in some cases reprocessing, of new science products developed by the research team.

Documentation - The Science Data Center generates complete documentation any new science products developed by the research team that constitute new research quality products to be made available to the

general science community (e.g. products cited in publications by members of the research team which should be available other scientists seeking to corroborate or extend the research performed by the team). The Science Data Center cooperates with a Backbone Data Center that receives its products after completion of its working life (or with the designated long term archive for its products) to ensure that documentation is brought to SEEDS adopted long term archiving standards.

Archive - The Science Data Center would not perform an archive function per se, but would maintain working storage of products obtained from other sources or science products generated as part of the research effort it supports.

Search and Order - The Science Data Center provides a search and order capability tailored to meet the needs of the research team it supports, perhaps supplemented to an additional capability for allowing other interested scientists to search for and order certain products the research team deems to be ready for use beyond the immediate work of the science team prior to their availability from a Backbone Data Center.

Access and Distribution - The Science Data Center will make the products collected to support the research effort readily available to members of the research team, and will perform reformatting, subsetting, or packaging of those products as needed to facilitate their interuse by the research team. The Science Data Center will also transfer new research quality science products and documentation to a Backbone Data Center when broader distribution of those products is appropriate, or at the conclusion of the research effort.

User Support - The Science Data Center provides close support to members of the research team it supports, including a help desk supplemented by on-line aids (e.g. FAQs).

Instrument / Mission Operations - None.

Sustaining Engineering - The Science Data Center performs software maintenance as needed.

Engineering Support - The Science Data Center performs engineering support functions as needed.

Technical Coordination - The Science Data Center participates in SEEDS system level processes, including coordination on data management, documentation standards, standards for content of life cycle data management plans, standards and best practices (including quality assurance standards and practices), interfaces, security, and common metrics.

The Science Data Center also participates in ESE level and/or bilateral processes to coordinate production and delivery of products between itself and other ESE data service providers.

Implementation - The Science Data Center develops the data and information system capabilities it requires by the to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

The Science Data Center also maintains an ongoing applications software development effort, developing 'science software' - product generation software embodying science algorithms, e.g. to produce research products to meet the needs of the research team and, in some cases, the overall ESE research program. In some cases the science software would be developed to run in the Science Data Center's own environment, in other cases the Science Data Center could provide science software to a Backbone Data Center for operational production.

Management - The Science Data Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Science Data Center includes planning information technology upgrades / technology refreshes, based on assessments of changing research team needs and availability of new technology. developing and maintaining life cycle data management plans (which address data migrations). It also includes coordinating its internal activities with the mission and science team it supports, and its

interaction with the ESE, collaboration among ESE data service providers to support science needs, and any other ESE or broader advisory activities that may be appropriate.

The Science Data Center also participates in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS.

Facility / Infrastructure - The Science Data Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and working storage for its data and information holdings. The Science Data Center provides a separate off-site backup of any new research quality science products generated by the research effort (e.g. that are cited by research team publications), for which the Science Data Center might use the services of a Backbone Data Center.

The Science Data Center ensures system and site security according to established NASA security policies and practices.

The Science Data Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

3.4   Systematic Measurements Centers

This section describes the generic Systematic Measurements Center type.

3.4.1   Systematic Measurements Center Concept

The following is the concept for NewDISS Multi-Mission Data Centers, herein referred to as Systematic Measurements Centers, from the NewDISS Concept Paper: "A fourth type of data center is the Multi-Mission Data Center. An example of the type of data activity to be carried out by such a data center is the generation of consistent time-series geophysical parameters, an activity exemplified by the current National Oceanic and Atmospheric Administration (NOAA)/NASA Pathfinder Datasets program, which is funded by NASA's ESE and carried out by PIs at various institutions. These efforts will take on more importance in the future, since NASA ESE has the requirement for generating time-series of geophysical parameters, while the EOS mission strategy has evolved so that it is now designed to accommodate technological change. Thus, these efforts will include construction of the long-time scale datasets from more than one NASA (or other) mission. These data centers will need to address network connectivity as part of their on-going activities. Selection of these services will be driven by PI-teaming arrangements, using either NASA-available resources or competitive alternatives."

3.4.2   Systematic Measurements Center Functions

In general, the Systematic Measurements Center is a potentially long lived data management capability implemented to support a particular data synthesis effort by a limited community of users (which will be called its 'synthesis team'). An example of a data synthesis effort would be research into how to cross-calibrate and consistently map measurements made by different missions (perhaps overlapping or consecutive) in order to be able to generate a consistent, continuous, long-term, research quality data set spanning multiple instruments/missions, validation of the cross-calibrated data sets, and then the production of the long time series data set. Such a production effort could be quite intensive in order to accomplish in a reasonable time the generation of a long time series data set involve handling many year's worth of a number of good sized data sets. The synthesis effort could continue adding new data sets to the mix from which its products are produced, extending its time series. The Systematic Measurements Center operates in a research environment, without the need for robustness and performance as would be the case for an operational environment.

The distinction drawn between a Science Data Center and a Systematic Measurements Center is that the former supports a particular research effort, while the latter supports a data synthesis effort that might require an extended period of time to complete, and which would enable future science efforts using the new, research quality, long time series data sets it produces.

Ingest - The Systematic Measurements Center obtains data and products and all supporting documentation needed for its data synthesis effort from a variety of sources, including other ESE data service providers, other agency data centers, etc. The ingest would not be performed on a time critical, operational basis, but could involve large amounts of data if long time series of large data sets are involved.

Processing - The Systematic Measurements Center would perform processing of new data synthesis products (such as long time series data sets) developed by the synthesis team on an ad hoc basis. This processing could be a major effort, for example if the objective is a long time series product produced from a number of large, multi-year input data sets. The Systematic Measurements Center could accomplish a large scale processing effort (such as a major effort to generate a long time series data set once the cross-calibration, mapping, etc., involved had been tested and validated) through a partnership with a Backbone Data Center or other processing facility.

Archive - The Systematic Measurements Center would not perform an archive function per se, but would maintain working storage of data and products obtained from other sources and new data synthesis products generated by the center. This could involve large data volumes, and the working storage would be configured to facilitate the processing effort.

Search and Order - The Systematic Measurements Center provides a search and order capability tailored to meet the needs of the synthesis team it supports, perhaps supplemented to an additional capability for allowing other interested scientists to search for and order certain products the synthesis team deems to be ready for use beyond the immediate work of the science team prior to their availability from a Backbone Data Center.

Access and Distribution - The Systematic Measurements Center generates complete documentation any new data synthesis products developed by the synthesis team that are new research quality products to be made available to the general science community, including full, documentation of the cross-calibration and any other steps taken to build the consistent time series. The Systematic Measurements Center will make the products collected to support the data synthesis effort readily available to members of the synthesis team. The Systematic Measurements Center will also transfer new research quality data synthesis products and documentation to a Backbone Data Center when broader distribution of those products is appropriate, or at the conclusion of the data synthesis effort.

User Support - The Systematic Measurements Center provides close support to members of the synthesis team it supports, including a help desk supplemented by on-line aids (e.g. FAQs).

Instrument / Mission Operations - None.

Sustaining Engineering - The Systematic Measurements Center performs software maintenance as needed.

Engineering Support - The Systematic Measurements Center performs engineering support functions as needed.

Technical Coordination - The Systematic Measurements Center participates in SEEDS system level processes, including coordination on data management, documentation standards, data stewardship (including standards for content of life cycle data management plans), standards and best practices (including quality assurance standards and practices), interfaces, security, common metrics, and interoperability (e.g. for data access and integration), across / within SEEDS and with other systems and networks as needed to support the ESE program.

The Systematic Measurements Center may cooperate with other ESE data service providers in representing ESE / SEEDS in broader community processes in areas such as standards, interoperability, data management, security, etc.

The Systematic Measurements Center also participates in SEEDS level and/or bilateral processes to coordinate production and delivery of products between itself and other ESE data service providers.

Implementation - The Systematic Measurements Center develops the data and information system capabilities it requires by the to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

The Systematic Measurements Center also maintains an ongoing applications software development effort. Applications software can include software to perform data services (e.g. subsetting, reformatting, reprojection, etc.) for more of its products, software tools for use by users to unpack, subset, or otherwise manipulate products provided by the Backbone Data Center, product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need, and to perform a 'data mining' or data integration operation to meet a user need.

Management - The Systematic Measurements Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Systematic Measurements Center includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. It includes developing data stewardship practices, performing data administration with science advice (via the User Advisory Group and other appropriate bodies), developing and maintaining life cycle data management plans (which address data migrations). It also includes coordinating its internal science activities and its interaction with the ESE and broader science community, including collaboration among ESE data service providers to support science needs, and any other ESE or broader advisory activities that may be appropriate.

The Systematic Measurements Center also participates in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS.

Facility / Infrastructure - The Systematic Measurements Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings, including a separate off-site backup facility for its new research quality data synthesis products, for which the Systematic Measurements Center might use the services of a Backbone Data Center.

The Systematic Measurements Center ensures system and site security according to established NASA security policies and practices.

The Systematic Measurements Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

3.5 Applications Center

This section describes the generic Applications Center.

3.5.1 Applications Center Concept

ESE's Applications Program mission ("Earth Science Enterprise Applications Strategy for 2002-2012", January 2002) is to "Expand and accelerate the realization of societal and economic benefits from Earth science, information, and technology". The overarching goal of the Applications Program is "to bridge the gap between Earth system science research results and the adoption of data and prediction capabilities for reliable and sustained use in decision support". Applications program implementation is seen as selecting applications projects (based on criteria discussed in the strategy document) that would proceed through the steps of applications research, validation and verification, and applications demonstration (depending on their maturity at startup). According to the strategy document, "the desired outcome of applications projects is for the partner organization to use the resulting prototypes, processes, and documentation as benchmarks for operational use. The desired impact is for the application to thrive because the service provider and it customers derive value

from the benefits of the operational use of Earth science in serving their decision making processes". At the conclusion of the project NASA would no longer be a source of funding.

Currently, over two hundred applications projects are ongoing (ESE Applications website), organized around several 'themes': resource management (over seventy projects currently), disaster management (over one hundred projects currently), community growth and infrastructure (over twenty projects currently), and environmental assessment (over twenty projects currently). The great majority of these applications projects are focused on developing solutions to specific problems or answers to specific questions within the 'theme' areas, and are not developing a data service provider capability. This prompted much discussion at the February, 2002, SEEDS workshop, which suggested that appropriate 'levels of service' would consider how well problems are solved or questions answered, how flexible the approaches taken were, how extensible the approaches taken would be outside of the specific context in which they were originally developed (e.g. would they work in different geographical regions, different social settings, etc.).

For the purposes of the LOS/CE study, only those current applications projects that can reasonably be seen as having at least some of the attributes of data service providers, will be considered as 'applications centers'. These are a small minority of the over two hundred applications projects ongoing. These will be taken as representative of functionally similar future activities, and information from these will be sought for the comparables database. Three types of applications activities can be viewed as current examples of 'applications centers' (although not every individual case within each type may be a good data service provider fit):

Regional Earth Science Applications Centers (RESACs)
The Regional Earth Science Applications Centers are designed to apply remote sensing and attending technologies to well-defined problems and issues of regional significance. There are currently nine public/private consortia throughout the U. S. that form seven RESACs. These consortia will apply state-of-the-art NASA Earth science research results to such diverse areas as precision farm management; monitoring of forest growth and health; regional water resources and hydrology; assessment of the impact of long-term climate variability and change; land cover and land use mapping; agricultural crop disease and infestation detection; management of fire hazards; watershed and coastal management; environmental monitoring; and primary and secondary science education.

Applications Earth Science Information Partners (Type 3 ESIPs)
The Earth Science Information Partners are drawn from academia, government and the private sector. Type 3 ESIPs are charged with developing innovative, practical applications of earth science data for the broader community. Eight Type 3 ESIPs are currently active.

Socio-Economic Data and Applications Center (SEDAC)

SEDAC, the Socioeconomic Data and Applications Center, is one of the Distributed Active Archive Centers (DAACs) in the Earth Observing System Data and Information System (EOSDIS) of the U.S. National Aeronautics and Space Administration. SEDAC focuses on human interactions in the environment. Its mission is to develop and operate applications that support the integration of socioeconomic and Earth science data and to serve as an "Information Gateway" between the Earth and social sciences.

The nine RESACs, eight applications ESIPs, and SEDAC are taken as current examples of Applications Centers, and even within this group there is considerable diversity in the size, scope, function of the activities which range from a NASA-funded DAAC to shared funding partnerships with private groups (universities, private corporations, etc.). SEDAC as a NASA funded EOSDIS DAAC is tightly coupled into EOSDIS and SEEDS processes. Type 3 ESIPs, members along with DAACs in the ESIP Federation, are currently involved in SEEDS processes, and at least while they receive some NASA funding can be expected to continue that involvement, and RESACs logically should be involved with SEEDS in the future. While applications groups receive NASA funding their participation in SEEDS and ESE processes can be supported, but once they become financially independent of NASA their continued participation would become their option.

It is possible to foresee a future in which, given NASA's commitment to applications expressed in the applications strategy document for 2002 - 2012, there will be a few enduring activities (e.g. SEDAC and some if not all RESACs) and a larger number of activities that arise from cooperative efforts (e.g. applications ESIPs) and go on to independence as they succeed. At any point in time there would be a mixture of enduring activities and other projects at different stages in their evolution, and a much larger number (e.g. about 200 currently) of focused applications projects at various stages in their work.

The intent of this section is to describe an Applications Center as that small subset of the full range of applications activities that have many or at least some of the attributes of a data services provider in that they are not simply targeted on a specific problem for a specific user but offer services or products to meet the needs of a broader community.

Applications Centers will obtain NASA Earth science products and use these, sometimes in conjunction with other Earth science data or any kind of other data to produce special products and/or deliver tailored services to an applications community. The products and services may be oriented around a particular problem or applications area. These communities could include agriculture, fisheries, urban planning, resource management, many etc., which could derive value from NASA Earth science products if they were suitably formatted or packaged (e.g. for use with Geographic Information System (GIS) technology) or used in conjunction with other data to produce new products specifically designed to meet their needs.

Although Applications Center type embraces a variety of possible functional models, the approach being taken to cost estimation, i.e. allowing the user of the cost estimation tool to pick needed functions from a general list, allows cost estimates for individual applications centers to be made. The use of an applications center type template would be less helpful in an individual case, but is expected to provide a reasonable approximation when an overall cost estimate for an ESE level data services architecture alternative is being examined.

3.5.2 Applications Center Functions

In general Applications Centers perform the same functions as other data service provider types, the primary distinction being the nature of their user community and therefore their products and services.

Ingest - The Applications Center obtains data and products required as inputs for its applications products from other ESE data service providers, other agency data centers, etc. In some cases the ingest would be performed on a time critical, operational basis, and in other cases might be on an ad hoc or intermittent basis, and could involve large amounts of data.

Processing - The Applications Center would perform processing of new applications products (such as products for agriculture or fisheries) developed by the Applications Center. This processing could be a major effort if low level data sets of large size are used to generate products on a routine basis. Data integration is likely to be a key processing task for Applications Centers, since use of combinations of ESE science products and a variety of different types of data (socio-economic, etc.) to produce new products is a central function.

Documentation - The Applications Center would generate documentation sufficient to support the current use of its products. Documentation would be written for the applications user, and might also include documentation according to standards in use in applications communities. In some cases, where applications products are to be retained for long term use, the Applications Center ensures that its products are documented to the SEEDS adopted standard for long term archiving, working as necessary with external sources (e.g. other data service providers) to capture all needed information. The Applications Center also ensures produces search metadata, product guides, etc., to SEEDS adopted standards, which would likely include FGDC compliant metadata.

Archive - The Applications Center would not be likely to perform an archive function per se, depending perhaps on the commercial value of its products beyond their first use, but would maintain working storage of data and products obtained from other sources and new applications products generated by the center. This

could involve large data volumes, and the working storage would be configured to facilitate the processing effort.

Search and Order - The Applications Center serves a broad user community with a robust and flexible search and order capability that supports user interaction with search and order services and, increasingly, supports automated search and order interaction between software running on a user system and the Applications Center system. The search capability allows an applications user to apply criteria that might include applications parameter(s), spatial-temporal coverage, specific product names, etc., to the metadata describing available products and returning to the user listings supplemented by descriptive information of those product types and instances that meet the criteria. The 'order' capability includes a request/permission step, regardless of how implemented (e.g. manual or automated), where a request for a set of data or product instances, perhaps the results of (or a selected subset of the results of) a search, is processed and accepted or denied.

Applications Center search and order includes providing local user interface and capability and may include providing an interface to a broader based, ESE cross-site search and order capability.

Access and Distribution - The Applications Center may distribute its products to either a very limited user community or a very broad user community, operationally or intermittently and/or on an request basis depending on its particular mission or business plan. A key data service provided by Applications Centers is to make its products readily useable by applications communities, for example by providing them in GIS formats, given the currently widespread and rapidly growing use of GIS tools by many applications groups.

User Support - The Applications Center provides effective user support for a focused or wide range of users depending on its particular mission. Its user support includes assistance provided in direct contact with users by user support staff, e.g. responding to queries, taking of orders, staffing a help desk (i.e., staff awaiting user contacts who can assist in ordering, track and status pending requests, resolve problems, etc.), etc. The Applications Center will increasingly offer more automated user support aids (beginning with on-line documentation, FAQ, etc.) to meet increasing demands on user support with the proliferation of data types, data sources, and tools for users. User support staff should include applications expertise to provide users with assistance in selecting and using data.

The Applications Center may coordinates its user support with other ESE data service providers (e.g. for user referral services). It performs outreach to potential new users, and may participate in coordinated outreach activities with other ESE data service providers.

Instrument / Mission Operations - None.

Sustaining Engineering - The Applications Center performs software maintenance as needed.

Engineering Support - The Applications Center performs engineering support functions as needed.

Technical Coordination - The Applications Center participates in SEEDS system level processes, including coordination on data management, documentation standards, data stewardship (including standards for content of life cycle data management plans), standards and best practices (including quality assurance standards and practices), interfaces, common metrics, and interoperability (e.g. for data access and integration), across / within SEEDS and with other systems and networks as needed to support the ESE program.

The Applications Center participates in ongoing examination of the changing needs of the ESE science and applications program and the consequent impacts on the roles, missions, and services of ESE data service providers.

The Applications Center might participate in coordination of user support guidelines and practices across the network of ESE data service providers and with other data centers as needed to support the ESE science and applications program.

The Applications Center cooperates with other ESE data service providers in representing ESE / SEEDS in broader community processes in areas such as standards, interoperability, data management, security, etc.

The Applications Center also participates in SEEDS level and/or bilateral processes to coordinate access and timely delivery of products its application effort requires from other ESE data service providers,

Implementation - The Applications Center develops the data and information system capabilities it requires by the to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

The Applications Center also maintains an ongoing applications software development effort. Applications software can include software to perform data services (e.g. subsetting, reformatting, reprojection, etc.) for more of its products, software tools for use by users to unpack, subset, or otherwise manipulate products provided by the Applications Center, product generation software embodying science algorithms, e.g. to produce an applications product to meet a particular user need, and to perform a 'data mining' or data integration operation to meet a user need.

Management - The Applications Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Applications Center includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. Depending on the nature of the applications activity, site-level management might include developing data stewardship practices, performing data administration with appropriate advice (via a User Advisory Group or other appropriate body), developing and maintaining life cycle data management plans (which address data migrations). It may also include coordinating its interaction with the ESE and broader applications and science community, collaboration among ESE data service providers to support applications and science needs, annual Enterprise peer review, and support for its User Advisory Group (which includes representation from the applications, education, etc., communities that it serves) and any other ESE or broader advisory activities that may be appropriate.

The Applications Center may also participate in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS.

Facility / Infrastructure - The Applications Center provided and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings, including a separate off-site backup facility for its data and information holdings. The Applications Center ensures system and site security according to appropriate security policies and practices, depending on its nature (e.g. commercial practices for private entities, established NASA security policies and practices for NASA funded entities).

The Applications Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

3.6   Information Center

This section describes the generic ESE Information Center type.

3.6.1   Information Center Concept

In general the Information Center performs many of the same functions as the Back Bone Data Center, except that the Information Center is concerned with information describing data and products (i.e., one or more types of metadata) rather than the data and products themselves. In general the Information Center will obtain its information from other data service providers, assemble it and make it available to its users, and when its users discover data or products they desire, then help (e.g. by providing links to data service provider websites) those users obtain access to the services of source data service providers.

The addition of this data service provider type was based on discussion at the Formulation Team Retreat, November 7-8, 2001, where 'Echo' was suggested as a possible future instance. The Global Change Master Directory (GCMD) is a currently operational instance of the Information Center type.

3.6.2 Information Center Functions

The paragraphs below will discuss the Information Center role in each of the general data service provider reference model's functional areas.

Ingest - The Information Center performs ingest of one or more metadata types, ranging from product instance (e.g. granule) level inventory metadata streams to overall product type descriptions or service descriptions. In some cases the ingest function may be performed on a time critical, operational basis, e.g. for inventory metadata received from other data service providers to be posted to the Information Center's inventory on an operational basis. In other cases, ingest of product type descriptions (etc.) are received on an ad hoc basis and are infrequently updated. Quality control on incoming metadata is critical if the Information Center's database is to be current with consistent and accurate content.

Processing - The Information Center does not perform this function.

Documentation - The Information Center ensures that is own content is consistent and complete and in conformance with adopted ESE / SEEDS standards, but does not generate or maintain any other documentation.

Archive - The Information Center provides working storage for its database of descriptive information.

Search and Order - The Information Center serves a broad user community with a robust and flexible search and order capability that supports user interaction with search and order services and, increasingly, supports automated search and order interaction between software running on a user system and the Information Center system. The search capability allows a user to apply criteria that might include geophysical parameter(s), spatial-temporal coverage, specific product names, etc., to the metadata describing available data and products and returning to the user listings supplemented by descriptive information of those data or product types that meet the user's criteria and, depending on the level of metadata held by the Information Center, data or product instances that meet specific criteria.

Information Center search and order can include providing local user interface and capability and/or providing an interface to a broader based, ESE cross-site search and order capability.

Access and Distribution - While it provides access and distribution of its own metadata, the Information Center facilitates access to the data and products its metadata describes. This might be in the form of links to source data service provider websites, or the ability to accept a user request for relay to a source data service provider.

User Support - The Information Center provides effective user support for a wide range of users who access its metadata holdings and to the source data service providers who provide the metadata.

Instrument / Mission Operations - None.

Sustaining Engineering - The Information Center performs sustaining engineering, with no or very infrequent interruption of operational capabilities.

Engineering Support - The Information Center performs engineering support functions with no or very infrequent interruption of its operations.

Technical Coordination - The Information Center participates in SEEDS system level processes, including coordination on documentation (especially metadata) standards, standards and best practices, interfaces, security, common metrics, and interoperability across / within SEEDS and with other systems and networks as needed to support the ESE program.

The Information Center participates in ongoing examination of the changing needs of the ESE science and applications program and the consequent impacts on the roles, missions, and services of ESE data service providers.

The Information Center participates in coordination of user support guidelines and practices across the network of ESE data service providers and with other data centers as needed to support the ESE science and applications program.

The Information Center cooperates with other ESE data service providers in representing ESE / SEEDS in broader community processes in areas such as standards, interoperability, data management, security, etc.

Implementation - The Information Center develops the data and information system capabilities it requires by the to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

Management - The Information Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

Site-level management by the Information Center includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. It includes performing data administration with science advice (via the User Advisory Group and other appropriate bodies), developing and maintaining a life cycle data management plan covering its information holdings). It also includes coordinating its interaction with the ESE and broader science community, collaboration among ESE data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group (which includes representation from the science, applications, education, etc., communities that it serves) and any other ESE or broader advisory activities that may be appropriate.

The Information Center also participates in ESE / SEEDS management processes, strategic planning, and coordination with other data centers and activities beyond ESE/SEEDS.

Facility / Infrastructure - The Information Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings, including a separate off-site backup facility for its metadata. The Information Center ensures system and site security according to established NASA security policies and practices.

The Information Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

The Information Center provides facility / infrastructure support at a level that ensures no or very infrequent interruption of its operations.

3.7   Long Term Archive Center

This section describes the generic Long Term Archive Center.

3.7.1 Long Term Archive Center Concept

The report, "Global Change Science Requirements for Long-Term Archiving" (USGCRP, March 1999), of the results of the science panel that met in a workshop held at the National Center for Atmospheric Research (NCAR) in October, 1998, discussed the essential functions and characteristics of a long term archiving program.

In general the Long Term Archive Center performs most if not all of the same functions as the Backbone Data Center, with the additional focus on permanent preservation and archiving of data and products and their documentation, and active support to climate research, etc., that requires reprocessing of and/or access to long

time series of data and products. The Long Term Archive Center participates with ESE data service providers in life cycle data management planning and in a process for obtaining science guidance and priorities for long term archiving.

Long term archiving is strictly speaking not an ESE responsibility, but inclusion of a hypothetical Long Term Archive Center type is intended to support planning that NASA is doing with NOAA and USGS, the agencies who have (with NARA) the long term archive responsibility.

3.7.2 Long Term Archive Center Functions

The paragraphs below will discuss the Long Term Archive Center role in each of the general data service provider reference model's functional areas, drawing on the USGCRP report cited in Section 3.7.1 above. Items in the functional discussion below that are explicitly derived from that report are indicated by an appended '(USGCRP)'.

Ingest - The Long Term Archive Center performs ingest of a wide variety of data and product types, ranging from low level data streams to ancillary data to all of the levels of derived products, and their documentation. These products may be new to the center or may be replacements of earlier versions of products already archived by the center.

It is essential that the Long Term Archive Center verify the integrity and quality of data and derived product and associated documentation as it is ingested into the archive (USGCRP).

The ingest would be a transfer from another data service provider, e.g. a Backbone Data Center, according to scenario to be documented in life cycle data management plans. If the transfer is from a research environment (e.g. a Science Data Center) that Long Term Archive Center should proactively reach out to the research source and develop the needed agreements and procedure, assist in planning documentation, etc., (USGCRP). The transfer could be a single bulk delivery, or staged as a series of deliveries over a period of time. The transfer could be by media or network.

Processing - It is essential that the Long Term Archive Center exercise data to produce new products and/or new versions of old products to validate data and product documentation, identify and resolve problems in the data, provide opportunities to scientists within the center to pursue science interests, produce new or updated products that are of value to the science community, provide an opportunity to rethink and reorganize how the data are stored to take into account user access needs as well as accommodate new storage and access technology, and increase data longevity (USGCRP). Typical science processing / reprocessing efforts could include production of long time series of intercalibrated data sets from multiple sources/ sensors to support climate change research.

Processing / reprocessing by the Long Term Archive Center would be on an ad hoc basis, but with tight quality control.

Documentation - It is most essential that the Long Term Archive Center ensure that its data sets and products in the archive are accompanied by complete, comprehensive, and accurate documentation (USGCRP), in accordance with long term archive documentation standard. The center works as necessary with external data sources (e.g. other data service providers) to capture all needed information.

Archive - The Long Term Archive Center provides a very robust archive capability, performing insertion of data into archive storage, and preservation of data, metadata, and documentation within the archive. Preservation and maintenance of data holdings, including ensuring integrity and quality of the data, products, and associated documentation is an essential function of the Long Term Archive Center (USGCRP). Extension of maintenance to include updating of documentation with user comments on the data or product is desirable (USGCRP).

Preservation measures should include quality screening of data entering and exiting the archive, quality screening of archive media, off-site backup with sampling to verify integrity, and accomplishing migrations from one type of media to another. It is essential that the Long Term Archive Center develop and maintain a

multi-year data migration plan, and that the center perform integrity checks on archive media between migrations (USGCRP).

Data migrations to new archive technology should be taken as opportunities for processing / reprocessing (USGCRP).

Search and Order - The Long Term Archive Center serves a broad user community with a robust and flexible search and order capability that supports user interaction with search and order services and, increasingly, supports automated search and order interaction between software running on a user system and the Long Term Archive Center system. The search capability allows a user to apply criteria that might include geophysical parameter(s), spatial-temporal coverage, specific product names, etc., to the metadata describing available data and products and returning to the user listings supplemented by descriptive information of those data or product types and instances that meet the criteria. The 'order' capability includes a request/permission step, regardless of how implemented (e.g. manual or automated), where a request for a set of data or product instances, perhaps the results of (or a selected subset of the results of) a search, is processed and accepted or denied.

Long Term Archive Center search and order can include providing local user interface and capability and/or providing an interface to a broader based, ESE cross-site search and order capability.

Access and Distribution - The Long Term Archive Center serves a broad user community with a robust search and order and distribution (electronic and media) service, including offering subsetting, reformatting, repackaging in response to user needs. It is essential that the center provide the next and subsequent generation of scientists with appropriate access to, and facilitate their use of, its holdings, where 'access' includes a data set / product search and order function, the ability to deliver data and/or products and supporting information (documentation) on suitable media or electronically, and choices of format, user options such as subsetting, that facilitate access and use (USGCRP).

User Support - The Long Term Archive Center provides effective user support (a user support staff knowledgeable about the data and products, willing and able to help users identify, obtain, and use the products the need, including making referrals to other sources of data - USGCRP) for a wide range of users.

Instrument / Mission Operations - None.

Sustaining Engineering - The Long Term Archive Center performs sustaining engineering, with no or very infrequent interruption of operational capabilities.

Engineering Support - The Long Term Archive Center performs engineering support functions with no or very infrequent interruption of its operations.

Technical Coordination - The Long Term Archive Center is participates in SEEDS system level processes, including coordination on data management, documentation standards, data stewardship (including standards for content of life cycle data management plans), standards and best practices (including quality assurance standards and practices), interfaces, common metrics, and interoperability (e.g. for data access and integration), across / within SEEDS and with other systems and networks as needed to support the ESE program.

The Long Term Archive Center participates in coordination of user support guidelines and practices across the network of ESE data service providers and with other data centers as needed to support the ESE science and applications program.

The Long Term Archive Center also participates in multi-lateral and/or bilateral processes to coordinate production and delivery of products between itself and other ESE data service providers.

Implementation - The Long Term Archive Center develops the data and information system capabilities it requires by the to perform its mission, including initial design and implementation of the data system (hardware and system software) and applications software and expansion or replacement (i.e. technology refresh) as needed over its operating life.

The Long Term Archive Center also maintains an ongoing applications software development effort. Applications software can include software to perform data services (e.g. subsetting, reformatting, reprojection, etc.) for more of its products, software tools for use by users to unpack, subset, or otherwise manipulate products provided by the Long Term Archive Center, product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need, and to perform a 'data mining' or data integration operation to meet a user need.

Management - The Long Term Archive Center performs a variety of site-level management functions as well as performing direct management of its functional areas.

provides management for its own operation and staff to support its participation in archive related activities. For example, it is essential that the center be actively facilitate the process for deciding which products to include or exclude from, or remove from, the archive (USGCRP). It is essential this process be driven by science priorities and scientific assessments, and that scientists be actively engaged in the process: setting criteria and making decisions (USGCRP). The Long Term Archive Center would participate with the appropriate ESE data service providers in these processes.

Site-level management by the Long Term Archive Center includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. It includes developing data stewardship practices, performing data administration with science advice, developing and maintaining life cycle data management plans (which address data migrations). It also includes coordinating its internal science activities and its interaction with the ESE and broader science community, including a visiting scientist program or equivalent, collaboration among ESE data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group (which includes representation from the science, applications, education, etc., communities that it serves) and any other ESE or broader advisory activities that may be appropriate.

Facility / Infrastructure - The Long Term Archive Center provides and maintains a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings, including a separate off-site backup facility for its data and information holdings. The Long Term Archive Center ensures system and site security according to established agency security policies and practices.

The Long Term Archive Center performs resource planning, logistics, supplies inventory and acquisition, and facility management. It provides for purchase of supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.'

The Long Term Archive Center provides facility / infrastructure support at a level that ensures no or very infrequent interruption of its operations.

4   Allocation of Requirements / LOS to Data Service Provider Types

This section presents the mapping of the general template of data service provider requirements and levels of service presented in Working Paper 5, "Data Service Provider Reference Model - Requirements / Levels of Service" to the ESE data service provider types discussed in Section 3 above. [The term 'template' is used because the requirements contain placeholders for specifics that must be filled in (i.e. choices between alternatives shown, or between possible levels of service, or replacement of placeholders with lists or numerical values) to generate from the template a set of requirements / levels of service that would apply to a specific ESE data service provider, and that would allow a cost estimate for it to be produced.] This mapping would be the basis for separate requirements / levels of service templates for each data service provider type. They in turn become the basis for the projection of estimated costs for new ESE data service providers of each type.

The requirements / levels of service templates will vary from data service provider type to data service provider type. The different types of data service provider will not all perform the same functions, and will not all meet the same requirements. Indeed, where different data service provider types do have a requirement in common, different levels of service are often appropriate for different data service provider types. The objective of the mapping is to show which of the general data service provider requirements apply to each data service provider type, and where applicable, to indicate minimum, recommended, and desirable levels of service for each requirement.

The tables in this section are arranged to allow convenient comparison of how the requirements / levels of service apply to the different data service provider types. The table below is a sample illustrating the format used in the tables below.

Sample Table

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall ingest…. (2.1 a)	1) operational (time-critical) ingest…	M	M			Dep	R	
	2) routine ingest and verification…					Dep	M	R
	3) ingest on a non-operational basis with verification…			R	R	Dep		M
	4) ingest on a non-operational basis.			M	M	Dep		

The first column in each table presents the requirement (with the section containing the requirement in Working Paper 5, "Data Service Provider Reference Model - Requirements / Levels of Service" in parentheses).

The second through ninth columns are an n by 7 matrix of n levels of service vs the seven logical data service provider types. The logical data set provider types are abbreviated:

> BBDC - Backbone Data Center
>
> MDC - Mission Data Center
>
> SDC - Science Data Center
>
> SMC - Systematic Measurements Center
>
> AC - Applications Center

IC - Information Center

LTAC - Long Term Archive Center

The second column contains the draft levels of service defined for the requirement, or "none" if there are no levels of service for the requirement. Entries in the next seven columns indicate if, and if so how, each level of service applies to each logical data service provider type. The possible entries for each cell are as follows:

N/a - the requirement does not apply to the logical provider type.

Blank - the requirement applies to the logical provider type, but the level of service does not.

M - the level of service is the minimum required of the provider type.

R - the level of service is recommended for the provider type.

D - the level of service is desired for the provider type.

Dep - for 'depends' - the requirement applies, but there is no predominant level of service for the provider type - real cases could be at any of the levels of service shown.

Y - when there are no levels of service and the requirement applies.

For a given data service provider type, the entries represent the predominant weight. For example, a data service provider type may ingest a number of different data streams, and a particular ingest level of service might apply for each one. What is indicated in this table is the ingest level of service that best characterizes the data service provider type, especially for the purpose of cost estimation. A similar example would be the backbone data service provider type, which might perform ad hoc as well as operational processing; in such a case the requirements / levels of service will reflect the operational processing. In any case, when a cost estimate is being made for an actual data service provider, its specific requirements would be used, so that in the previous example its cost estimate would not reflect operational processing if it's mission did not include any.

The "desired" case can arise for Applications Centers, which may receive NASA/ESE funding only temporarily, or Long Term Archive Centers which are funded by their host agency.

Note that minimum and recommended levels of service may be indicated, or minimum, recommended and desirable levels of service.

There are a few cases where an actual data service provider of given type might not meet a particular requirement contrary to what is indicated in the table. For example, if a Science Data Service Provider provides its data and products to a Backbone Data Center, then the requirement under distribution calling for a data service provider to provide its data, products, and documentation to a Long Term Archive Center would not apply to that data service provider, and a cost estimate for that Science Data Service Provider would reflect that.

As indicated above the mapping in these tables would be used to write a set of requirements / levels of service templates, one for each ESE data service provider type. Each template could then be turned into a high level requirements statement for a specific data service provider of its type by filling the items left as placeholders in the template.

The next several pages contain the requirements / levels of service to data service provider type mappings.

4.1 Ingest

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall ingest the following data [ingest data stream table, listing for each data stream: name, source, product types ingested, product type format, products ingested per day of each type, volume ingested per day]. The input data streams should cover all data to be received by the center, e.g. satellite data streams, ancillary data products, processed products generated by other data service providers, etc., based on its ESE mission. (2.1 a)	1) operational (time-critical) ingest with immediate verification of data integrity and quality;	M	M			Dep	R	
	2) routine ingest and verification of data quality and integrity without tight time constraints;					Dep	M	R
	3) ad hoc or intermittent ingest on a non-operational basis with verification of data quality and integrity;			R	R	Dep		M
	4) ad hoc or intermittent ingest on a non-operational basis.			M	M	Dep		

Ingest levels of service can be mixed within a data service provider; i.e. different levels may be appropriate for different data streams.

Ingest requirements for Applications Centers can vary widely from case to case.

4.2 Processing

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall generate the following standard products, included required Level 1B products [standard product table, listing for each product type/series: name, format, retention plan, product instances produced per day, volume per day, required input data streams] on a highly reliable, operational basis, either on a routine schedule or on-demand, based on its ESE mission. (2.2 a)	1) standard products shall be generated within 2 days of ingest/availability of required inputs.	D	D	n/a	n/a	Dep	n/a	n/a
	2) standard products shall be generated within 7 days of ingest/availability of required inputs.	R	R	n/a	n/a	Dep	n/a	n/a
	3) standard products shall be generated within 30 days of ingest / availability of required inputs.	M	M	n/a	n/a	Dep	n/a	n/a
The data service provider shall generate the following products [product table, listing for each product type/series: name, format, retention plan, average product instances produced per day, average volume per day, required input data streams] on an ad hoc, non-operational basis. (2.2 b)	1) specific targets for processing adopted on a case by case basis.	Y	n/a	R	D	Dep	n/a	D
	2) general goals for processing.	Y	n/a	M	R	Dep	n/a	R
	3) no goals, purely ad hoc processing.	Y	n/a		M	Dep	n/a	M
The data service provider shall reprocess standard products [standard product table] on an ad hoc basis in response to reprocessing requests. (2.2 c)	1) the capacity for reprocessing shall be 9 times the original processing rate.	D	D			n/a	n/a	R
	2) the capacity for reprocessing shall be 6 times the original processing rate.	R	R			n/a	n/a	M

	3) the capacity for reprocessing shall be 3 times the original processing rate.	M	M	R	R	n/a	n/a	
The data service provider shall reprocess standard products [standard product table, listing for each product a reprocessing interval] according to a reprocessing schedule. (2.2 d)	1) reprocessing shall be performed according to a negotiated reprocessing schedule.	R	R			n/a	n/a	R
	2) reprocessing shall be performed to meet the general goals of a nominal schedule.	M	M	R	R	n/a	n/a	M
	3) reprocessing shall be performed following a nominal schedule on a resource / time available basis.			M	M	n/a	n/a	

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall accept science algorithm software from users for [product list], and perform integration and test of the software, and operational execution of the software to produce products. (2.2 e)	1) the data service provider shall accept standard, research product generation software, and/or data integration and data mining software from users;	D	n/a	n/a	n/a	n/a	n/a	
	2) the data service provider shall accept research product generation software and/or data integration and data mining software from users;	R	n/a	n/a	n/a	n/a	n/a	R
	3) the data service provider shall accept standard and/or research product generation software from users;	M	n/a	n/a	n/a	n/a	n/a	
	4) the data service provider shall accept research product generation software from users;		n/a	n/a	n/a	n/a	n/a	M
	5) the data service provider shall accept standard product generation software from users.		n/a	n/a	n/a	n/a	n/a	
The data services provider shall be capable of cross-calibration of data from multiple sources to produce consistent product time series spanning multiple instruments / platforms. (2.2 f)	None.	D	n/a	n/a	Y	n/a	n/a	D
The data service provider shall provide standard metrics on production to the SEEDS Office. (2.2 g)	None.	Y	Y	Y	Y	n/a	n/a	n/a

The processing level of service can vary for different product generation tasks within a site.

Science Data Center and Systematic Measurements Centers would accept, integrate, test, and execute science software developed by their research teams, but (it is assumed) not from other users.

4.3   Documentation

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall generate and provide ESE/SEEDS standard compliant catalog information (metadata, including browse) and documentation describing all data and information held by the data service provider. (2.3 a)	1) data and product holdings (including multiple versions of products and corresponding documentation as needed) documented to the ESE / SEEDS adopted standard for long term archiving, including details of processing algorithms, processing history, many etc.	M	R	R	M		n/a	M
	2) documentation ensured to be sufficient for current use (e.g. product type descriptions, product instance (a.k.a. granule) descriptions including version information, FAQs, 'readme's, web pages with links to metadata, user guides, references to journal articles describing the production or use of the data or product).		M	M		R	n/a	
	3) documentation only as received from product provider.					M	n/a	
The data service provider shall update documentation of data and products with user comments. (2.3 b)	1) data and products routinely updated with user comments.	D					n/a	R
	2) data and products occasionally updated with user comments.	R					n/a	M
	3) data and products rarely updated with user products.	M	R	R	R	R	n/a	
The data service provider shall generate and provide DIF (directory interchange format) documents to the Global Change Master Directory on all products available from the data service provider prior to their re-lease for distribution. (2.3 c)	None	Y	Y	Y	Y	Dep	n/a	Y

4.4 Archive

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall add to its archive or working storage the following data and products [archive product table, drawn from ingest data stream table, standard product, and ad hoc product tables and reprocessing volume] and related documentation / metadata. (2.4 a)	None	Y	Y	Y	Y	Dep	n/a	Y
The data service provider shall provide for secure, permanent storage of data at the "raw" sensor level (NASA Level 0 plus appended calibration and geolocation information). (2.4 b)	None	Y	Y	n/a	n/a	n/a	n/a	Y
The data service provider shall provide for secure storage of all standard or other science products it produces until the end of the science mission or until transfer to an approved permanent archive, per the applicable life cycle data management plan (or separate retention plan). (2.4 c)	None	Y	Y	Y	Y	n/a	n/a	n/a
The data service provider shall have the capability to selectively replace archived product instances (single or large sets) with new versions, and to selectively update metadata and documentation (e.g. to update quality flags when a product is validated). (2.4 d)	None	Y	n/a	Y	Y	n/a	Y	Y
The data service provider shall provide for an [archive] [working storage] capacity of [number] TB. (2.4 e)	1) archive capacity is cumulative sum of all data ingested plus all products generated (including allowance for retaining multiple versions of the same product as required to provide needed support to the provider's science or applications community).	M	M	R	R	Dep	n/a	M
	2) archive capacity is limited to a specified threshold.			M	M			
The data service provider shall perform quality screening on data entering the archive (e.g. read after write check when data is written to archive media) and exiting the archive (e.g. track read failures and corrected errors or other indication of media degradation on all reads from archive media). (2.4 f)	1) exit and entry screening.	R	n/a	n/a	n/a	n/a	n/a	M
	2) entry screening.	M	n/a	n/a	n/a	n/a	n/a	

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall take steps to ensure the preservation of data in its archive. (2.4 g)	1) 10% per year random screening.	D	n/a	n/a	n/a	n/a	n/a	R
	2) 5% per year random screening.	R						M
	3) 1% per year random screening.	M						
The data service provider shall provide a backup for its [archive] [working storage]. (2.4 h)	1) full off-site backup, with regular sampling to verify integrity.	M	R			n/a		M
	2) partial, [Backup Fraction - % of archive backed up], off-site backup, with sampling.		M	R	R	n/a		
	3) partial, [Backup Fraction - % of archive backed up], on-site backup with sampling.			M	M	n/a	M	
The data service provider shall use robust archive media. (2.4 i)	1) archive media compliant with best commercial practice.	M				n/a	n/a	M
	2) archive media and system vendor independent.		R	R	R	n/a	n/a	
	3) archive media vendor independent.		M	M	M	n/a	n/a	
The data service provider shall plan and perform periodic migration of archive to new archive media / technology. (2.4 j)	1) planned migration.	R	n/a	n/a	R	n/a	n/a	M
	2) no planned migration, but ad hoc migration as need is seen to arise.	M	n/a	n/a	M	n/a	n/a	
The data service provider shall provide standard metrics on archive to the SEEDS Office. (2.4 k)	None	Y	n/a	n/a	Y	Y	Y	D

4.5 Search and Order

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall provide users with access to all metadata and information holdings. (2.5 a)	1) public access to all users.	M					M	M
	2) access to the science and applications community.			R	R	M		
	3) access to a limited team of scientists.		M	M	M			
The data service provider shall provide a world wide web accessible search and order capability to [all users (including the general public) consistent with SEEDS standards and practices][to a limited set of science team members]. (2.5 b)	1) allow search for instances of multiple product types that pertain to a specified object or phenomenon (e.g. a named hurricane, a volcanic eruption, a field campaign, etc.).	D				n/a	D	D
	2) allow search for instances of multiple product types by geophysical parameter(s), time, and space applied across multiple product types.	R				n/a	R	R
	3) allow search for instances of multiple product types by common time and space criteria (coincident search).	M	D	D	D	n/a	M	M
	4) allow search for instances of single product type by time and space criteria.		R	R	R	n/a		
	5) allow search for particular instances of a product type from a list of those available.		M	M	M	n/a		
The data service provider shall provide the user with the option of quickly viewing information describing any product returned as meeting search criteria. (2.5 c)	1) descriptive information includes detailed algorithm and use explanations, references to a few published papers that describe the production or use of the product, standard guide and DIF metadata.	D	D	D	D	n/a	n/a	D
	2) descriptive information includes references to a few published papers that describe the production or use of the product, standard guide and DIF metadata.	R	R	R	R	n/a	n/a	R
	3) descriptive information includes standard guide and DIF metadata.	M	M	M	M	n/a	n/a	M
The data service provider shall provide an interface for system-system search and order access as well as an interface for human users. (2.5 d)	None.	Y	n/a	n/a	n/a	Dep	Y	Y
The data service provider shall provide an interface to and support selected external catalog search capabilities. (2.5 e)	None.	Y	n/a	R	R	Dep	n/a	Y

Search and order requirements and levels of service for the Science Data Center and Systematic Measurements Center that go beyond meeting the needs of the science teams they support apply when these providers retain science quality products and make them more widely available.

4.6 Access and Distribution

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall provide users with access to all data and product holdings, including all standard science products (Level 1b, Level 2, and Level 3) produced by the data service provider. (2.6 a)	1) public access to all users.	M				Dep	M	M
	2) access to the science community.			R	R	Dep		
	3) access to a limited team of scientists.		M	M	M	Dep		
The data service provider shall provide data and products to users in (at a minimum) one of the SEEDS core formats. (2.6 b)	None.	Y	Y	Y	Y	n/a	n/a	Y
The data service provider shall enhance its distribution capability with supporting data services such as subsetting, resampling, reformatting (e.g. to GIS formats), reprojection and/or packaging to meet the needs of its users. (2.6 c)	1) supporting data services available for most archived data and products.	R				Dep	n/a	R
	2) supporting data services available for less than half of archived data and products.	M	R	R	R	Dep	n/a	M
	3) supporting data services available for a few selected data and products only.		M	M	M	Dep	n/a	
The data service provider shall provide data to users on an [operational, subscription, and/or in response to request] basis. (2.6 d)	None.	Y	Y	Y	Y	Y	n/a	Y
The data service provider shall provide an interface for system to system network delivery of data and products. (2.6 e)	None.	Y	R	D	D	D	n/a	Y
The data service provider shall perform timely distribution of data and products to users by network, providing an average distribution volume capacity of [number] TB per day. (2.6 f)	1) availability of a single product for access by user software within ten seconds.	D				Dep	n/a	D
	2) availability of a single product for network delivery within ten seconds.	R	D	D	D	Dep	n/a	R
	3) availability of a single product for network delivery within ten minutes.	M	R	R	R	Dep	n/a	M
	4) availability of a single product for network delivery within twenty four hours.		M	M	M	Dep	n/a	

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall perform timely distribution of data and products to users on SEEDS standard media types in response to user requests, providing an average volume capacity of [number] TB per day. (2.6 g)	1) shipping of media product within three days of receipt of request.	R				Dep	n/a	D
	2) shipping of media product within one week of receipt of request.	M	R	R	R	Dep	n/a	R
	3) shipping of media product within one month of receipt of request.		M	M	M	Dep	n/a	M
The data service provider shall have the capacity to distribute products on an average of [number] media units per day. (2.6 h)	None.	Y	Y	Y	Y	Y	n/a	Y
The data service provider with final ESE archive responsibility (i.e., a Backbone Data Center unless, for example, a Science Data Service Provider held its products to the time for their transfer to the long term archive) shall transfer its data, products, and documentation (done to the long term archive standard) to the designated long term archive according to its Life Cycle Data Management Plan. (2.6 i)	None	Y	Y	Y	Y	n/a	n/a	n/a
The data service provider shall provide SEEDS standard metrics on distribution to the SEEDS Office. (2.6 j)	None	Y	Y	Y	Y	Y	n/a	n/a

4.7 User Support

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall be capable of supporting [number] of distinct, active users per year who request and use data service provider products. (2.7 a)	1) one user support staff member per 100 active users.	R	M	M	R	Dep		
	2) one user support staff member per 500 active users.	M			M	Dep		R
	3) one user support staff member per 1,000 active users.						M	M
The data service provider shall provide a trained user support staff. (2.7 b)	1) below plus science expertise in data / product quality and their research uses.	R	R	R	R	Dep	n/a	R
	2) below plus technical expertise in data structures, use of tools for format conversions, subsetting, analysis, etc.	M				Dep	n/a	M
	3) below plus comprehensive knowledge of details of formats for most if not all products.		M	M	M	Dep	n/a	
	4) user support staff are knowledgeable about the data service provider's holdings and ordering/delivery options.					Dep	n/a	
The data service provider shall provide a help desk function (i.e., staff awaiting user contacts who can assist in ordering, track and status pending requests, resolve problems, etc.). (2.7 c)	1) Help desk staffed seven days per week, twenty-four hours per day.	R				Dep	R	R
	2) Help desk staffed five days per week, twelve hours per day;	M	R	R	R	Dep	M	M
	3) Help desk staffed five days per week, eight hours per day;		M	M	M	Dep		
The data service provider shall provide on-line user support (FAQ, data / product and service descriptions, etc.). (2.7 d)	None	Y	Y	Y	Y	Y	Y	Y
The data service provider shall perform user outreach, education, and training. (2.7 e)	1) Below plus provide user training sessions at universities, schools, etc.	D	n/a	n/a	n/a	Dep	D	D
	2) Below plus expanded booth support including mini-workshops, user training sessions;	R	n/a	n/a	n/a	Dep	R	R
	3) Below plus booth support at four conferences per year;	M	n/a	n/a	n/a	Dep	M	M
	4) Produce and make available outreach material - pamphlets, brochures, posters, etc.		n/a	n/a	n/a	Dep		

4.8 Instrument / Mission Operations

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC

Requirement		BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall monitor the status and performance of [name] instruments and in some cases also [name] spacecraft for which it is responsible, generating instrument commands and in some cases space-craft commands as needed. (2.8 a)	None.	n/a	Y	n/a	n/a	n/a	n/a	n/a
The data service provider shall obtain the services of a NASA (or other spacecraft operator as appropriate) mission operations facility to provide instrument and spacecraft data and to receive, validate, and transmit instrument and/or spacecraft commands to the spacecraft. (2.8 b)	None.	n/a	Y	n/a	n/a	n/a	n/a	n/a

4.9   Sustaining Engineering

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall maintain and, as needed, enhance custom software it develops to meet its mission needs, and reused software it customizes and integrates, a total of [number] SLOC. (2.9 a)	1) no or very infrequent interruptions of data service provider operations.	R	R				R	R
	2) occasional interruptions in data service provider operations.	M	M	R	R	R	M	M
	3) as needed, with interruptions in data service provider operations a secondary concern.			M	M	M		

4.10 Engineering Support

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall perform system administration, network administration, database administration, coordination of hardware maintenance by vendors, and other technical functions as required for performance of its mission. (2.10 a)	1) no or very infrequent interruptions of data service provider operations.	R	R				R	R
	2) occasional interruptions in data service provider operations.	M	M	R	R	R	M	M
	3) as needed, with interruptions in data service provider operations a secondary concern.			M	M	M		
The data service provider shall perform systems engineering, test engineering, configuration management, COTS procurement, installation of COTS upgrades, network / communications engineering and other engineering functions as required for performance of its mission. (2.10 b)	1) no or very infrequent interruptions of data service provider operations.	R	R				R	R
	2) occasional interruptions in data service provider operations.	M	M	R	R	R	M	M
	3) as needed, with interruptions in data service provider operations a secondary concern.			M	M	M		

4.11 Technical Coordination

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC

Requirement		C1	C2	C3	C4	C5	C6	C7
The data service provider shall provide staff required for participation in SEEDS processes, including ESE data services architecture refinement and evolution, and information technology planning. (2.11 a)	None.	Y	n/a	n/a	n/a	Y	Y	n/a
The data service provider shall provide staff required for participation in SEEDS processes to coordinate data stewardship standards and practices and development and maintenance of standards for content of life cycle data management plans. (2.11 b)	None.	Y	n/a	Y	Y	Y	n/a	Y
The data service provider shall provide staff required for participation in SEEDS processes to coordinate best practices among ESE data service providers, including quality assurance standards and practices for all phases of data services provider functions. (2.11 c)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide staff required for participation in SEEDS processes, and cooperating with other ESE data service providers in representing ESE / SEEDS in broader community processes, for developing and maintaining common standards and interface definitions, including those that enable interoperability within the ESE / SEEDS environment and with other systems and networks as needed to support the ESE program. (2.11 d)	None.	Y	Y	Y	Y	Y	Y	Y
The data services provider shall participate in SEEDS level and/or bilateral processes to coordinate production and delivery of products between ESE data service providers. (2.11 e)	None.	Y	Y	Y	Y	Y	Y	Y
The data services provider shall participate in SEEDS processes for coordinating user support guidelines and practices among ESE data services providers. (2.11 f)	None.	Y	n/a	n/a	n/a	Y	Y	Y
The data services provider shall provide staff required for SEEDS coordination of security standards and practices to meet NASA or other established security requirements. (2.11 g)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide staff to coordinate standards for common metrics. (2.11 h)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide funding for travel to support technical coordination activities. (2.11 i)	None.	Y	Y	Y	Y	Y	Y	Y

Participation by Applications Centers in SEEDS technical coordination would be expected to vary from case to case, depending on the specific mission of each one, and in any given year the degree to which it receives NASA funding. Participation of Applications Centers that become self-sustaining would depend on their view of the benefits of participation.

Long Term Archive Centers are presumed to be funded and operated by other agencies than NASA, i.e. NOAA and USGS. Their participation in SEEDS technical coordination would depend on agreements with NASA and/or their view of the benefits of participation.

4.12 Implementation

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall design and a data and information system capable of meeting its mission requirements. The design shall address hardware configuration and interfaces and allocation of function to platform. The design shall address software configuration, including COTS, software re-use, and new custom software to be developed, including science software embodying product generation algorithms and/or software facilitating integration of science software provided by outside source(s). (2.12 a)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall develop a staffing plan that addresses staff required to implement and operate the data service provider over its planned lifetime. The staffing plan shall include a breakdown of positions and skill levels assigned to functions. (2.12 b)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall develop a facility plan, including planning for space, utilities, furnishings, etc., required to support its staff, data and information system, data storage, etc., and the environmental conditioning to be provided. (2.12 c)	None.	Y	Y	Y	Y	Y	Y	Y
The data service provider shall accomplish the implementation of its data and information system, including purchase and installation of hardware, purchase or licensing and installation and configuration of COTS software, modification, installation and configuration of re-use software, development of new custom software, and integration of all components into a tested system capable of meeting the data service provider's mission requirements. (2.12.d)	None.	Y	Y	Y	Y	Y	Y	Y

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall perform ongoing applications software development. (2.12 e)	1) Below plus implementation of applications software to perform a 'data mining' or data integration operation to meet a user need.	D				D	n/a	D
	2) Below plus implementation of product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need.	R	M	M	M	R		R
	3) Implementation of software tools for use by users to unpack, subset, or otherwise manipulate products provided by the data service provider.	M				M		M
The data service provider shall provide the staff needed to accomplish all needed in-house development and test activities. (2.12 f)	None.	Y	Y	Y	Y	Y	Y	Y

4.13 Management

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
The data service provider shall provide management and administrative staff to perform supervisory, financial administration, and other administrative functions. (2.13 a)	None	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide staff required for participation in SEEDS management processes, strategic planning, coordination with other data centers and activities beyond ESE/SEEDS. (2.13 b)	None	Y	Y	Y	Y	D	Y	D
The data service provider shall provide staff with science expertise to coordinate the science activities within the data service provider and its interaction with the ESE and broader science community, including a visiting scientist program (or equivalent) , collaboration among ESE data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group and any other advisory activities appropriate given its ESE role and user community. (2.13 c)	None	Y	n/a	n/a	n/a	Dep	n/a	Y
The data service provider shall provide staff with system engineering expertise to plan information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. (Coordination with other ESE data service providers is included in technical coordination). (2.13 d)	None	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide staff with data management expertise to develop data stewardship practices, perform data administration with science advice (via the User Advisory Group and other appropriate bodies), develop and maintain life cycle data management plans including data migrations. (Coordination with other ESE data service providers is included in technical coordination). (2.13 e)	None	Y	n/a	n/a	n/a	Dep	n/a	Y

4.14 Facility / Infrastructure

Requirement	Levels of Service	BBDC	MDC	SDC	SMC	AC	IC	LTAC
2.14 a: The data service provider shall maintain site, system, and data security according to established NASA or other policies and practices while providing easiest possible access (consistent with required security) to its data and information services for its user community.	None	Y	Y	Y	Y	Y	Y	Y
2.14 b: The data service provider shall provide and maintain a fully furnished and equipped, environmentally con-trolled, physically secure facility to house its staff, systems, and data and information holdings.	None	Y	Y	Y	Y	Y	Y	Y
The data service provider shall provide a backup facility for its data and information holdings. (2.14 c)	1) an environmentally controlled and physically secure off-site backup archive facility.	R	R	D	D			M
	2) an on-site but separate environmentally controlled and physically secure off-site backup facility.	M	M	R	R	R	R	
	3) a backup capability within the data service provider's primary data system(s).			M	M	M	M	
The data service provider shall perform resource planning, logistics, supplies inventory and acquisition, and facility management. (2.14 d)	1) no or very infrequent interruptions of data service provider operations.	M	M	D	D	Dep	M	M
	2) occasional interruptions in data service provider operations.			R	R			
	3) as needed, with interruptions in data service provider operations a secondary concern.			M	M			
The data service provider shall provide network connections and services as needed to support its operations. (2.14 e)	None.	Y	Y	Y	Y	Y	Y	Y

References and Acronym List

The References Section and the Acronym List for all of these Working Papers is in the document "References and Acronyms for the Levels of Service / Cost Estimation Working Papers ".